SPSS Level III

Analysis for Categorical Data in Medical Research

Audrey S. Chang, Ph.D.
Biometrics Section
Clinical Study Service
Department of Clinical Investigation
WRAMC

8/13/04

SPSS Level III Outline

1. **Kappa Statistics**
2. **Cochran Q Statistics**
3. **McNemar Statistics**
4. **Cochran and Mantel-Haenszel Statistics**

1. Kappa Statistics

- **Validity (Accuracy)**
  - **agreement with Known Gold Standard**
- **Intra-Rater Reliability (or Agreement)**
  - **for reproducibility**
- **Inter-Rater Reliability (or Agreement)**
  - **for generalizability**

Calculation

$$Kappa = \frac{(\pi_0 - \pi_e)}{(1 - \pi_e)}$$

π_0 = Observed agreement

π_e = Agreement due to chance

Hypothesis testing

Ho : Kappa = 0

Observed agreement is no better than the agreement due to chance.

Ha : Kappa $\neq$ 0

Observed agreement exceeds that due to chance

Interpretation

Kappa = 1
Kappa = 0
Kappa > .75
Kappa = .40 to .75
Kappa < .40

Example #1 Thoracic Pedicle Screw Study

- 108 thoracic pedicle screw tracts prepared from 54 cadaveric thoracic vertebrate:

  - 43 with no violations

  - 65 with violations in medial (31), lateral (23), and anterior (11)

- Three surgeons were blinded
  - independently read the scans
  - repeated readings at 3 different time points

Exercise #1.1- Validity or Accuracy

- To determine how well a Surgeon can
  - 1) correctly detect the normal and abnormal pedicle screw tracts and
  - 2) correctly identify the location of the violations
- Agreement with known gold standard

SPSS Commands

Open data from S:\Stat_Class\Kappa.sav

Analyze

Descriptive Statistics

Crosstabs

For Row – Enter rater_a1, rater_a2, rater_a3

For Column – Enter key

For Statistics – Check Kappa

For Cells – Check Observed, Expected, and Total

Exercise #1.1- Validity or Accuracy

- **Observed Probability**
  = .389 + .167 + .148 +.093
  =.797
- **Expected Probability**
  = (.398)(.546) + (.287)(.185) +
  (.213)(.157) + (.102)(.111)
  = .315
- Kappa = (.797 - .315) / (1 - .315)
  = .703

Output Interpretation

- **Diagonal – total observed agreement**
- **Expected agreement by chance**

- **Kappa for the Validity (Accuracy) of Rater A at time 1 =.702**

- **Conclusion**

  P-value =.000 (or P < .0005)

Exercise #1.2 Intra-Rater Agreement

- To determine how well the Surgeon can reproduce his/her ratings over 3 different time points

SPSS Commands

Analyze

Descriptive Statistics

Crosstabs

For Row(s) – Enter rater_a1

For Column(s) – Enter rater_a2 and rater_a3

For Statistics – Check Kappa

For Cells – Check Observed, Expected, and Total

Output Interpretation

- **Diagonal – total observed agreement**
- **Expected agreement by chance**
- **Kappa for the Reliability between Time 1 and Time 2 = .745**

  P-value = .000 (or < .0005)
- **Kappa for the Reliability between Time 1 and Time 3 = .763**

  P-value= .000

Exercise #1.3 Inter-Rater Agreement

- To determine the agreement between two raters

SPSS Commands

Analyze

Descriptive Statistics

Crosstabs

For Row(s) – Enter rater_a1

For Column(s) – Enter rater_b1

For Statistics – Check Kappa

For Cells – Check Observed, Expected, and Total

2. Cochran Q Statistics

- To analyze repeated categorical response data with only two categories (yes/no) over several occasions or time points

- Example - A clinician might evaluate patients at weekly intervals regarding whether a new drug treatment is successful

Exercise 2

- 43 patients with a chronic condition.
- Treated by each of the three drugs.
- Assume no carry-over effects.
- Outcome measure – favorable or unfavorable reaction (0=no response; 1=response)

Hypothesis testing

Ho : Three drugs have the same response rate

Ha : The response rates are different

SPSS Commands

Open data from
S:\Stat_Class\CochranQ.sav

Analyze

Nonparametric Tests

K-related samples…

For Test variables Drug A, Drug B, Drug C

Check Cochran’s Q and undo Friedman

Click Statistics -> check Descriptive

Interpretation

Descriptive statistics: Response rate

value 0 = no response; 1 = response

Drug A = 0.61 (28/46)

Drug B = 0.61 (28/46)

Drug C = 0.35 (16/46)

- **P-value =.014**
- **Conclusion: The response rates are statistically significant different among the 3 drugs**

McNemar Test

- To compare difference between 2 related proportions
- Outcome is binary (yes/no)
- Measurements were repeated on the same subjects at
  - 2 different occasions, or
  - 2 time points

Exercise 3

- **Use the data from exercise 2**
- **To conduct *post hoc* pair-wise comparisons – to determine which two drugs are statistical different**

SPSS Commands

Analyze

Nonparametric Tests

2-related samples…

Move druga – drugb to the test pairs list druga – drugc

drugb - drugc

Check McNemars and uncheck Wilcoxon

Click Options -> check Descriptive

Interpretation

Descriptive statistics - Response rate

- **P-value**

  P = 1.0 for Drug A vs. B

  P = .038* for Drug A vs. C

  P = .038* for Drug B vs. C

- *** Drug A and B are statistically significant different from Drug C**

4. Cochran Mantel-Haenszel

- To determine the common odds ratio of two treatment groups over several study centers while controlling for the confounding effect by the center differences

- Treatment outcome is binary (response/no-response)

Exercise 4

- Data from Walker (1997)
- A multicenter study - 4 centers
- Subjects were randomized to dermotel (active treatment) or sodium hyaluronate (placebo)
- To compare the odds ratio between the active treatment group vs. the control group in the cure of the disease

SPSS Commands

- Open data S:\Stat_Class\Level III data\mh.sav
- Analyze
  - Descriptive Statistics
  - Crosstabs
  - For Row – Enter trt
  - For Column – Enter resp
  - For Statistics – Check Risk, Exact, and Cochran's and Mantel-Haenszel statistics
  - For Cells – Check Observed, Row, and Column

Output

- Response rate in each treatment group was computed for each center
- Note the Sample size in each center
- Fisher Exact test – compared treatment effect at each center
- Odds ratio estimated for each center
- Homogeneity test – to compare if the odds ratio are same across all centers
- Conditional Independence test

Interpretation

- **The odds ratios between dermotel group and the placebo group in percentage of cure are not statistically significant across the 4 centers. (P=.595 by Breslow-Day)**
- **The common odds ratio (OR) is 2.147 (P=0.041 by Cochran's test and P=0.068 by M-H test)**
- **The 95% CI for the common odds ratio is 1.019 to 4.520, ie the true OR lies somewhere between 1.02 and 4.52.**